

# DESIGN OF SYNTHETIC DNA PROBES.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15														
A.	M	-	N	-	P	-	N	-	N	-	R	-	S	-	E	-	H	-	D	-	T	-	I	-	K	-	T	-	T
	ATG	AAT	CCN	AAT	AAT	CGN	TCN	GAA	CAT	GAT	ACN	ATT	AAA	ACN	ACN														
		C			C	C	AGA	AGT	G	C	C		C	G															
							G	C						A															

B. ATGAATCCTAATAATCG  
 C C C C  
 A  
 G

C. GAACATGATACAATTAA  
 G C C G C  
 A

- A. PROTEIN SEQUENCE OF THE N-TERMINI OF PEAKS A AND B OF THE B.T.T. TOXIN AND DEDUCED DNA SEQUENCE.
- B. SYNTHETIC A1 PROBE, 32-FOLD DEGENERATE 17-MER, BASED ON AMINO ACIDS 1-6.
- C. SYNTHETIC A2 PROBE, 48-FOLD DEGENERATE 17-MER, BASED ON AMINO ACIDS 8-13.

FIG.1

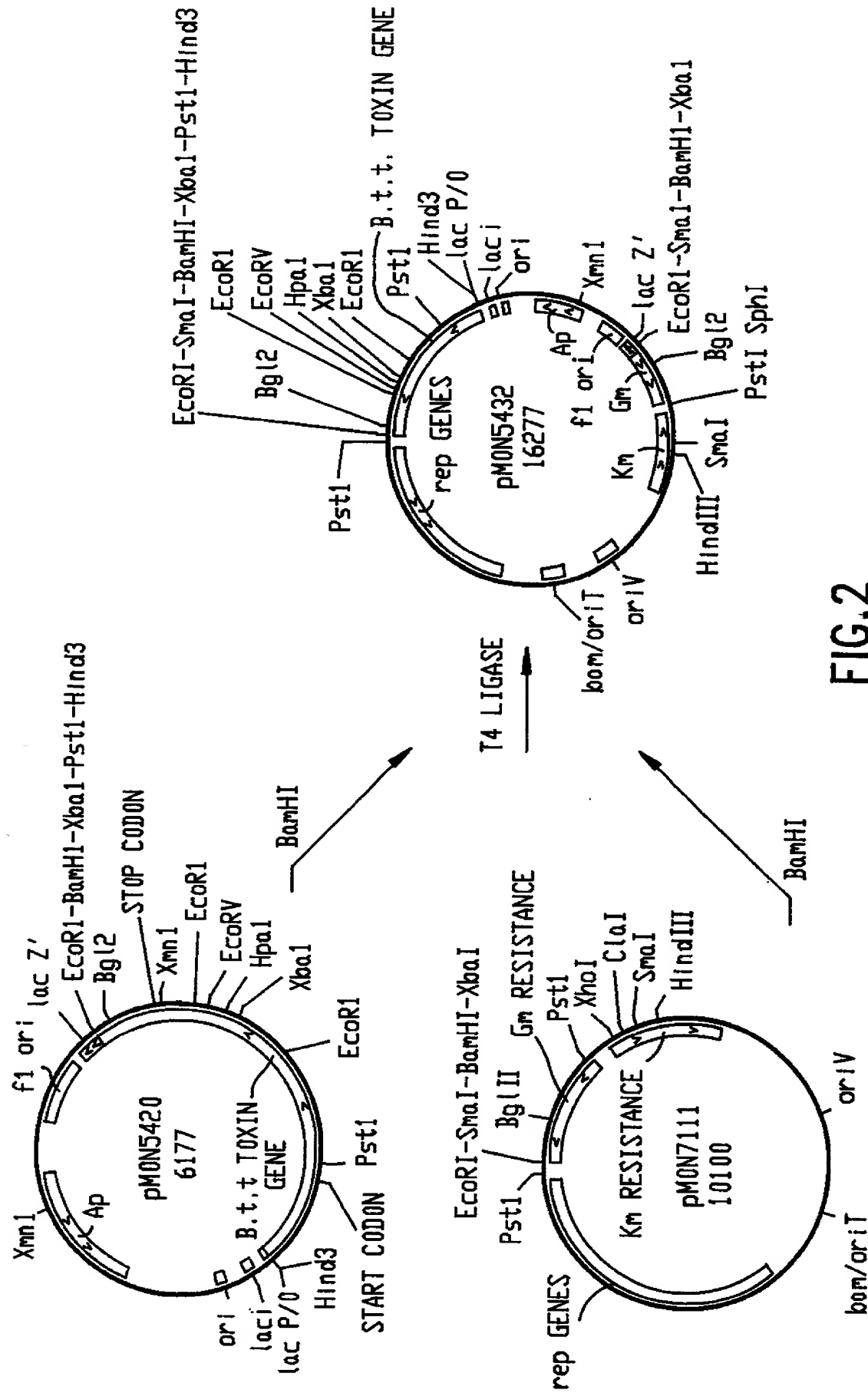


FIG.2

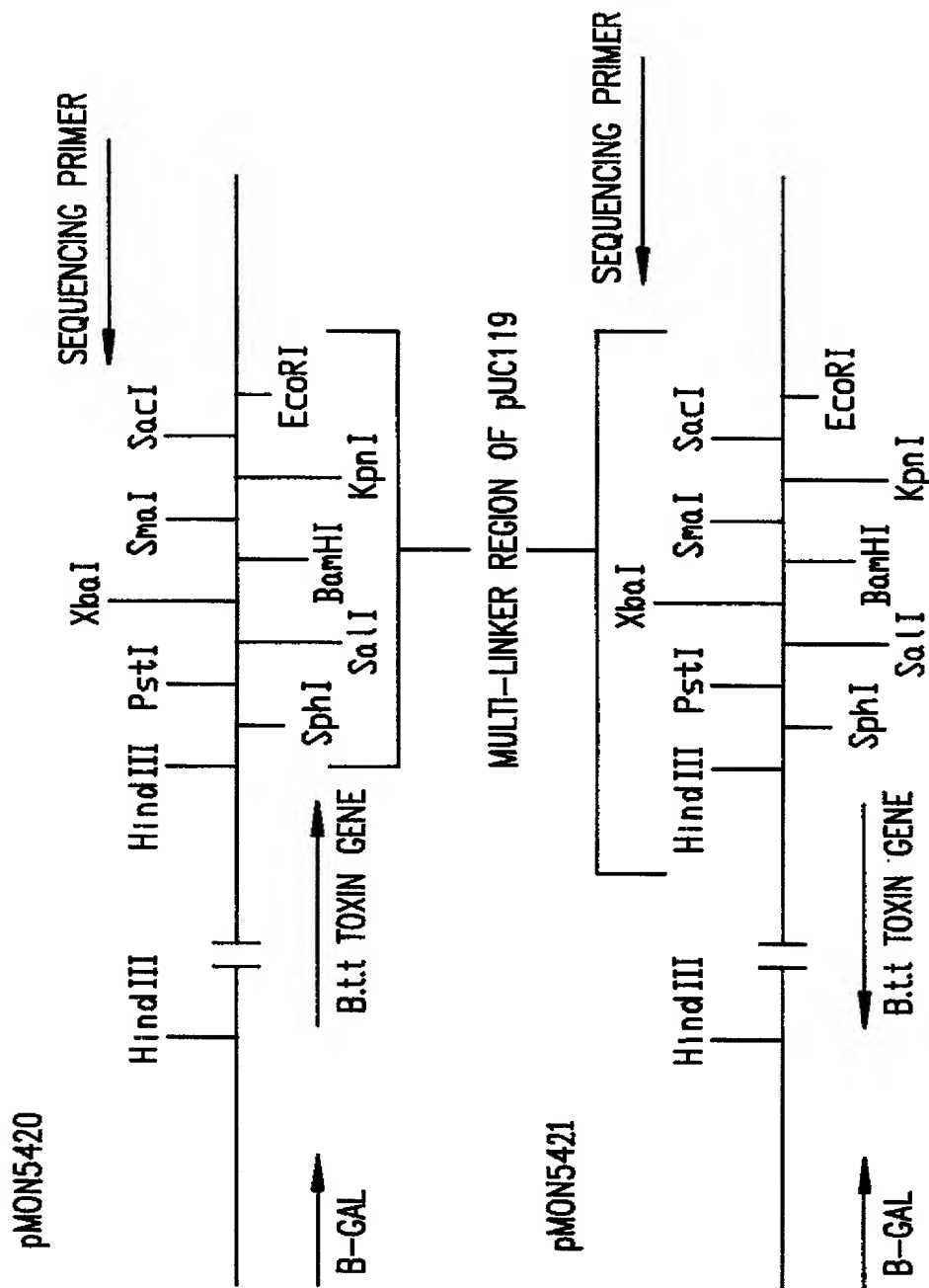


FIG.3

# SEQUENCING OF THE *Bacillus Thuringiensis* var. *Tenebrionis*

## INSECTICIDAL TOXIN GENE

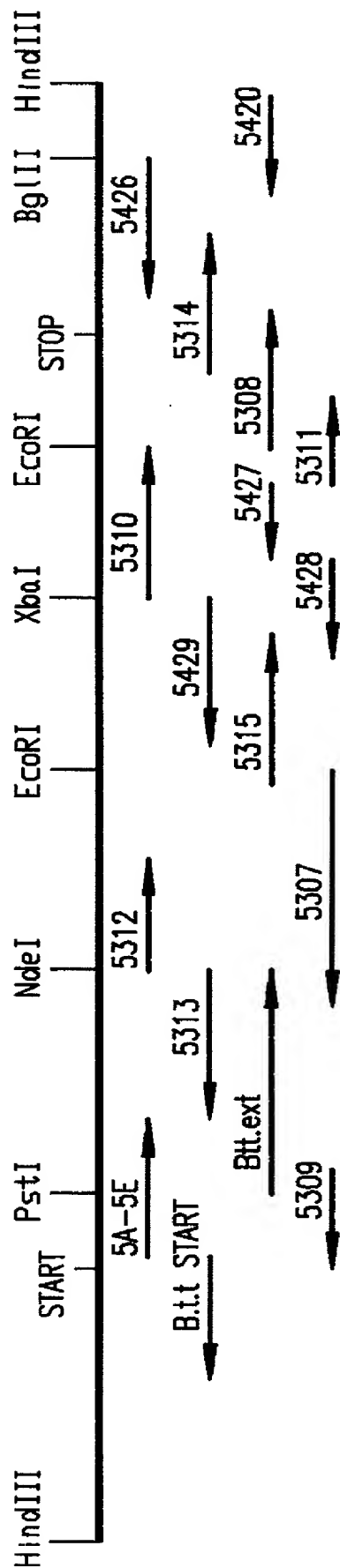


FIG.4

# SEQUENCE OF THE B.t.t. INSECTICIDAL TOXIN GENE AND FLANKING REGIONS

H  
i  
n  
f  
1

gagcgactattataatcatatcatatctTATTGGAATGATTAAGATTCCAATAGAATAG  
1 -----+-----+-----+-----+-----+-----+ 60  
ctcgtgataatattagtatgtataaaagATAACCTTACTAATTCTAAGGTTATCTTATC

S  
M f F M  
n a o b  
l N k o  
1 1 1 2

TGTATAAATTATTTATCTTGAAAGGAGGGATGCCTAAAAACGAAGAACATTAAAAACATA  
61 -----+-----+-----+-----+-----+-----+ 120  
ACATATTTAATAAATAGAACTTTCCTCCCTACGGATTTTGTCTTCTTGTAAATTTTGTAT

TATTTGCACCGTCTAATGGATTTATGAAAAATCATTTTATCAGTTTGAAAATTATGTATT  
121 -----+-----+-----+-----+-----+-----+ 180  
ATAAACGTGGCAGATTACCTAAATACTTTTGTAGTAAATAGTCAAACCTTTTAATACATAA

H  
M i M T N  
n n b a l  
l f o q a  
1 1 2 1 3

ATGATAAGAAAGGGAGGAAGAAAAATGAATCCGAACAATCGAAGTGAACATGATACAATA  
181 -----+-----+-----+-----+-----+-----+ 240  
TACTATTCTTTCCCTCCTTCTTTTACTTAGGCTTGTTAGCTTCACTTGTAATGTAT

START M N P N N R S E H D T I -

FIG.5A

	M	M		B	N		N
	a	n		a	l		l
	e	l		n	a		a
	2	1		1	4		3

AAAACTACTGAAAATAATGAGGTGCCAACTAACCATGTTCAATATCCTTTAGCGGAACT  
 241 -----+-----+-----+-----+-----+-----+ 300  
 TTTTGATGACTTTTATTACTCCACGGTTGATTGGTACAAGTTATAGGAAATCGCCTTTGA  
  
 K T T E N N E V P T N H V Q Y P L A E T -

	M		D	M		M	P
	a		r	b		a	s
	e		a	o		e	t
	1		1	2		2	1

CCAAATCCAACACTAGAAAGATTTAAATTATAAAGAGTTTTTAAGAATGACTGCAGATAAT  
 301 -----+-----+-----+-----+-----+-----+ 360  
 GGTTTAGGTTGTGATCTTCTAAATTTAATATTTCTCAAAAATTCTTACTGACGTCTATTA  
  
 P N P T L E D L N Y K E F L R M T A D N -

	M	A
	a	l
	e	u
	1	1

AATACGGAAGCACTAGATAGCTCTACAACAAAAGATGTCATTCAAAAAGGCATTTCCGTA  
 361 -----+-----+-----+-----+-----+-----+ 420  
 TTATGCCTTCGTGATCTATCGAGATGTTGTTTTCTACAGTAAGTTTTTCCGTAAAGGCAT  
  
 N T E A L D S S T T K D V I Q K G I S V -

FIG.5B

**THE UNIVERSITY OF CHICAGO**

/

481 -----+-----+-----+-----+-----+ 540  
TGT TTGAAAAATTTATGATAAACCGGTTCACTTCTGGGCACCTTCCGAAAATACCTTGT  
  
T N F L N T I W P S E D P W K A F M E Q -

GTAGAAGCATTTGATGGATCAGAAAATAGCTGATTATGCAAAAAATAAAGCTCTTGCAGAG  
541 -----+-----+-----+-----+-----+-----+ 600  
CATCTTCGTAACCTAGTCTTTTATCGACTAATACGTTTTTTATTTTGAGAACGTCTC  
  
V E A L M D Q K I A D Y A K N K A L A E -

**FIG.5C**

S			
DqH	T	M	N
rua	a	b	l
a9e	q	o	a
263	1	2	3
/			

TTACAGGGCCTTCAAAATAATGTCGAAGATTATGTGAGTGCATTGAGTTCATGGCAAAA  
 601 -----+-----+-----+-----+-----+ 660  
 AATGTCCCGGAAGTTTTATTACAGCTTCTAATACTCACGTAAGTCAAGTACCGTTTT  
  
 L Q G L Q N N V E D Y V S A L S S W Q K -

	BS	
	sc	A
	tr	l
	NF	u
	11	1
	/	

AATCCTGTGAGTTCACGAAATCCACATAGCCAGGGCGGATAAGAGAGCTGTTTTCTCAA  
 661 -----+-----+-----+-----+-----+ 720  
 TTAGGACACTCAAGTGCTTTAGGTGTATCGGTCCCCGCCCTATTCTCTCGACAAAAGAGTT  
  
 N P V S S R N P H S Q G R I R E L F S Q -

	M
	n
	l
	1

GCAGAAAGTCATTTTCGTAATTCATGCCTTCGTTTGCAATTTCTGGATACGAGGTTCTA  
 721 -----+-----+-----+-----+-----+ 780  
 CGTCTTTCAGTAAAAGCATTAAGTTACGGAAGCAAACGTTAAAGACCTATGCTCCAAGAT  
  
 A E S H F R N S M P S F A I S G Y E V L -

FIG.5D



			F	
	B	N	An	
	b	d	Lu	
	v	e	u4	
	1	1	1H	

TTTCTAACACATATGCACAAGCTGCCAACACACATTTATTTTACTAAAAGACGCTCAA  
 781 -----+-----+-----+-----+-----+-----+-----+ 840  
 AAAGATTGTTGTATACGTGTTTCGACGGTTGTGTGTAAATAAAAAATGATTTTCTGCGAGTT  
  
 F L T T Y A Q A A N T H L F L L K D A Q -  
  

H		M		M
g		b		b
a		o		o
1		2		2

 ATTTATGGAGAAGAATGGGGATACGAAAAAGAAGATATTGCTGAATTTTATAAAAGACAA  
 841 -----+-----+-----+-----+-----+-----+-----+ 900  
 TAAATACCTCTTCTTACCCCTATGCTTTTCTTCTATAACGACTTAAATATTTTCTGTT  
  
 I Y G E E W G Y E K E D I A E F Y K R Q -  
  

		M		M
		a		n
		e		i
		2		1

 CTAAACTTACGCAAGAATATACTGACCATTGTGTCAAATGGTATAATGTTGGATTAGAT  
 901 -----+-----+-----+-----+-----+-----+-----+ 960  
 GATTTTGAATGCGTTCTTATATGACTGGTAACACAGTTTACCATATTACAACCTAATCTA  
  
 L K L T Q E Y T D H C V K W Y N V G L D -  
  

		H	
		i	
		n	
		f	
		1	

 AAATTAAGAGGTTTCATCTTATGAATCTTGGGTAACTTTAACCGTTATCGCAGAGAGATG  
 961 -----+-----+-----+-----+-----+-----+-----+ 1020  
 TTTAATTCTCCAAGTAGAATACTTAGAACCCATTTGAAATTGGCAATAGCGTCTCTCTAC  
  
 K L R G S S Y E S W V N F N R Y R R E M -

FIG.5E

ACATTAACAGTATTAGATTTAATTGCACTATTTCCATTGTATGATGTTCCGGCTATACCCA  
 1021 -----+-----+-----+-----+-----+-----+ 1080  
 TGTAAATTGTCATAATCTAAATTAACGTGATAAAGGTAACATACTACAAGCCGATATGGGT  
  
 T L T V L D L I A L F P L Y D V R L Y P -  
  
 S H H  
 aX i i  
 uh n n  
 3o f c  
 A2 1 2  
 /  
 AAAGAAGTTAAAACCGAATTAACAAGAGACGTTTAAACAGATCCAATTGTCGGAGTCAAC  
 1081 -----+-----+-----+-----+-----+-----+ 1140  
 TTTCTTCAATTTTGGCTTAATTGTTCTCTGCAAAATTGTCTAGGTTAACAGCCTCAGTTG  
  
 K E V K T E L T R D V L T D P I V G V N -  
  
 DM AT.  
 ds sa  
 et uq  
 12 21  
 / /  
 AACCTTAGGGGCTATGGAACAACTTCTCTAATATAGAAAATTATATTGAAAACCAT  
 1141 -----+-----+-----+-----+-----+-----+ 1200  
 TTGGAATCCCCGATACCTTGTGGAAGAGATTATATCTTTTAATATAAGCTTTTGGTGTA  
  
 N L R G Y G T T F S N I E N Y I R K P H -  
  
 E BS  
 c T N sc  
 o h l tr  
 R a a NF  
 1 1 4 11  
 /  
 CTATTTGACTATCTGCATAGAATTCAATTTACACGCGGTTCCAACCAGGATATTATGGA  
 1201 -----+-----+-----+-----+-----+-----+ 1260  
 GATAAACTGATAGACGTATCTTAAGTTAAAGTGTGCGCCAAGGTTGGTCTATAATACCT  
  
 L F D Y L H R I Q F H T R F Q P G Y Y G -

FIG.5F

H	S	S
i	Aa H	M
n	vu p	a
f	a9 a	e
1	26 2	1
	/	

AATGACTCTTTCAATTATTGGTCCGGTAATTATGTTTCAACTAGACCAAGCATAGGATCA  
 1261 -----+-----+-----+-----+-----+-----+ 1320  
 TTACTGAGAAAGTTAATAACCAGGCCATTAATACAAAGTTGATCTGGTTCGTATCCTAGT

N D S F N Y W S G N Y V S T R P S I G S -

E	P	
c	f	R
o	l	S
P	M	a
1	1	1

AATGATATAATCACATCTCCATTCTATGGAAATAAATCCAGTGAACCTGTACAAAATTTA  
 1321 -----+-----+-----+-----+-----+-----+ 1380  
 TTACTATATTAGTGTAGAGGTAAGATACCTTTATTTAGGTCACTTGGACATGTTTTAAAT

N D I I T S P F Y G N K S S E P V Q N L -

E	
c	H
o	a
P	e
1	3

GAATTTAATGGAGAAAAAGTCTATAGAGCCGTAGCAAATACAAATCTTGCGGTCTGGCCG  
 1381 -----+-----+-----+-----+-----+-----+ 1440  
 CTTAAATTACCTCTTTTTTCAGATATCTCGGCATCGTTTATGTTTAGAACGCCAGACCGGC

E F N G E K V Y R A V A N T N L A V W P -

FIG.5G

M S  
a Ba  
e cu  
3 13  
1A  
/  
TCCGCTGTATATTCAGGTGTTACAAAAGTGAATTTAGCCAATATAATGATCAAACAGAT  
1441 -----+-----+-----+-----+-----+ 1500  
AGGCGACATATAAGTCCACAATGTTTTACCTTAAATCGGTTATATTACTAGTTTGTCTA  
S A V Y S G V T K V E F S Q Y N D Q T D -  
R H  
R i HT AP i  
s s n hh lv n  
a a f aa uu f  
1 1 1 11 12 1  
/  
GAAGCAAGTACACAAACGTACGACTCAAAAAGAAATGTTGGCGCGGTCAGCTGGGATTCT  
1501 -----+-----+-----+-----+-----+ 1560  
CTTCGTTCAATGTTTGCATGCTGAGTTTTCTTTACAACCGCGCCAGTCGACCCCTAAGA  
E A S T Q T Y D S K R N V G A V S W D S -  
S  
CaT M XM M N  
lua n ba n l  
a3q l ae l a  
1A1 1 11 1 3  
/  
ATCGATCAATTGCCTCCAGAAACAACAGATGAACCTCTAGAAAAGGGATATAGCCATCAA  
1561 -----+-----+-----+-----+-----+ 1620  
TAGCTAGTTAACGGAGGTCTTTGTTGTCTACTTGGAGATCTTTTCCCTATATCGGTAGTT  
I D Q L P P E T T D E P L E K G Y S H Q -

FIG. 5H

H  
iH  
np  
ca  
21  
/  
CTCAATTATGTAATGTGCTTTTTAATGCAGGGTAGTAGAGGAACAATCCCAGTGTTAACT  
1621 -----+-----+-----+-----+-----+-----+ 1680  
GAGTAAACATTACACGAAAAATTACGTCCCATCATCTCCTTGTTAGGGTCACAATTGA  
  
L N Y V M C F L M Q G S R G T I P V L T -  
  
H  
N i AT  
l n sa  
a f uq  
3 1 21  
/  
TGGACACATAAAAGTGTAGACTTTTTAACATGATTGATTGAAAAAAATTACACA ACTT  
1681 -----+-----+-----+-----+-----+-----+ 1740  
ACCTGTGTATTTTCACATCTGAAAAAATTGTACTAACTAAGCTTTTTTAAATGTGTTGAA  
  
W T H K S V D F F N M I D S K K I T Q L -  
  
B P S  
M s ADpAMaS M  
a p vruvaut n  
e M aaMr9y l  
3 1 2212161 1  
// //

CCGTTAGTAAAGGCATATAAGTTACAATCTGGTGCTTCCGTTGTGCGAGGT CCTAGGTTT  
1741 -----+-----+-----+-----+-----+-----+ 1800  
GGCAATCATTTCGTATATTCAATGTTAGACCACGAAGGCCAACAGCGTCCAGGATCCAAA  
  
P L V K A Y K L Q S G A S V V A G P R E -

FIG. 5 I

	E		F	
	c		n	M
	o		u	a
	R		4	e
	V		H	3

ACAGGAGGAGATATCATTCAATGCACAGAAAATGGAAGTGGGCAACTATTTACGTTACA

1801 -----+-----+-----+-----+-----+ 1860

TGTCCTCCTCTATAGTAAGTTACGTGTCTTTACCTTCACGCCGTTGATAAATGCAATGT

T G G D I I Q C T E N G S A A T I Y V T -

				E	
H	R F	T	AM	c	D
p	s o	a	la	o	d
a	a k	q	ue	R	e
2	1 1	1	11	1	1

CCGGATGTGTCGTACTCTCAAAAATATCGAGCTAGAATTCATTATGCTTCTACATCTCAG

1861 -----+-----+-----+-----+-----+ 1920

GGCCTACACAGCATGAGAGTTTTATAGCTCGATCTTAAGTAATACGAAGATGTAGAGTC

P D V S Y S Q K Y R A R I H Y A S T S Q -

		B	
D	B Ns	T	
d	a lp	a	
e	n a1	q	
1	1 42	1	

ATAACATTTACACTCAGTTTAGACGGGGCACCATTTAATCAATACTATTTGATAAAACG

1921 -----+-----+-----+-----+-----+ 1980

TATTGTAAATGTGAGTCAAATCTGCCCCGTGGTAAATTAGTTATGATAAAGCTATTTTGC

I T F T L S L D G A P F N Q Y Y F D K T -

ATAAATAAAGGAGACACATTAACGTATAATTCATTTAATTTAGCAAGTTTCAGCACACCA

1981 -----+-----+-----+-----+-----+ 2040

TATTTATTTCTCTGTGTAATTGCATATTAAGTAAATTAATCGTTCAAAGTCGTGTGGT

I N K G D T L T Y N S F N L A S F S T P -

FIG.5J

AT	H	AM
sa	g	ha
uq	a	ae
21	1	23
/		

2041 TTCTGAATTATCAGGGAATAACTTACAAATAGGCGTCACAGGATTAAGTGCTGGAGATAAA  
 -----+-----+-----+-----+-----+-----+ 2100  
 AAGCTTAATAGTCCCTTATTGAATGTTTATCCGCAGTGTCTAATTCACGACCTCTATTT  
  
 F E L S G N N L Q I G V T G L S A G D K -

	X	M
	m	a
	n	e
	1	1

2101 GTTTATATAGACAAAATTGAATTTATTCCAGTGAATTAATTAAGTAAAGTAAAGAAG  
 -----+-----+-----+-----+-----+-----+ 2160  
 CAAATATATCTGTTTTAACTTAAATAAGGTCACCTAATTTAATTGATCTTTCATTTCTTC  
  
 V Y I D K I E F I P V N x STOP

M	M
a	b
e	o
3	2

2161 TAGTGACCATCTATGATAGTAAGCAAAGGATAAAAAATGAGTTCATAAAATGAATAACA  
 -----+-----+-----+-----+-----+-----+ 2220  
 ATCACTGGTAGATACTATCATTCGTTTCCTATTTTTTTACTCAAGTATTTTACTTATTGT

	M
	b
	o
	2

2221 TAGTGTTCTTCAACTTTTCGCTTTTTGAAGGTAGATGAAGAACACTATTTTTATTTTCAAA  
 -----+-----+-----+-----+-----+-----+ 2280  
 ATCACAAGAAGTTGAAAGCGAAAACTCCATCTACTTCTTGTGATAAAAAATAAAAGTTT

FIG.5K

	D		D	
	r		r	
	a		a	
	1		1	

2281 ATGAAGGAAGTTTTAAATATGTAATCATTTAAAGGGAACAATGAAAGTAGGAAATAAGTC  
 -----+-----+-----+-----+-----+-----+-----+ 2340  
 TACTTCCTTCAAAATTTATACATTAGTAAATTTCCCTTGTTACTTTCATCCTTTATTCAG

			S	
			s	
			p	
			1	

2341 ATTATCTATAACAAAATAACCATTTTTATATAGCCAGAAATGAATTATAATATTAATCTT  
 -----+-----+-----+-----+-----+-----+-----+ 2400  
 TAATAGATATTGTTTTATTGGTAAAAATATATCGGTCTTTACTTAATATTATAATTAGAA

				H	
		A	D	iH	S
		l	d	ng	s
		u	e	fa	p
		1	1	11	1

2401 TTCTAAATTGACG†TTTTCTaAACGTTCTATAGCTTCAAGACGCTTAGAATCATCAATAT  
 -----+-----+-----+-----+-----+-----+-----+ 2460  
 AAGATTTAACTGCaAAAAGA†TTGCAAGATATCGAAGTTCTGCGAATCTTAGTAGTTATA

				H	
A	A	T		i	B
c	l	a		n	g
c	u	q		f	l
1	1	1		1	2

2461 TTGTATACAGAGCTGTTGTTTCCATCGAGTTATGTCCCATTTGATTGCTAATAGAACAA  
 -----+-----+-----+-----+-----+-----+-----+ 2520  
 AACATATGTCTCGACAACAAAGGTAGCTCAATACAGGGTAAACTAAGCGATTATCTTGTT

FIG.5L



S			
aX		M	F
uh		n	o
3o		l	k
A2		1	1
//			

GATCTTTATTTTCGTTATAATGATTGGTTGCATAAGTATGGCGTAATTTATGAGGGCTTT

2521 -----+-----+-----+-----+-----+-----+ 2580

CTAGAAATAAAAGCAATATTACTAACCAACGTATTCATACCGCATTAAATACTCCCGAAA

TCTTTTCATCCAAAAGCCAAGTGTATTTCTCTGTA

2581 -----+-----+-----+----- 2615

AGAAAAGTAGGTTTTTCGGTTCACATAAAGAGACAT

FIG.5M

B.t.t. PROTEINS – NATIVE & <u>E.coli</u> CLONES					
BAND No.	MW (Kd)	NATIVE B.t.t.	E.coli CLONES		
			pMON5436	pMON5456 pMON5450	pMON5460
1	73	—	—		—
2	71	—			
3,3'	67	—	—	—	
4	66	—			
DIAGRAM OF B.t.t. PROTEINS. B.t.t. PROTEINS PRODUCED BY <i>Bacillus Thuringiensis</i> var. <i>Tenebrionis</i> and <i>E.coli</i> JM101 (pMON5436, pMON5456, pMON5450, pMON5460) WERE RESOLVED ON 9% SDS-PAGE AND THE RESPECTIVE PATTERNS ARE SHOWN.					

FIG.6

## FIG. 7

N-TERMINI OF THE UNIQUE B.t.t. PROTEINS PRODUCED IN EITHER B.t.t. AND/OR E. COLI WERE DETERMINED BY AMINO ACID SEQUENCING. THE ARROWS AND ASSOCIATED NUMBERS CORRESPOND TO THE FIRST AMINO ACID OF THE PROTEINS DESCRIBED IN FIG. 6.

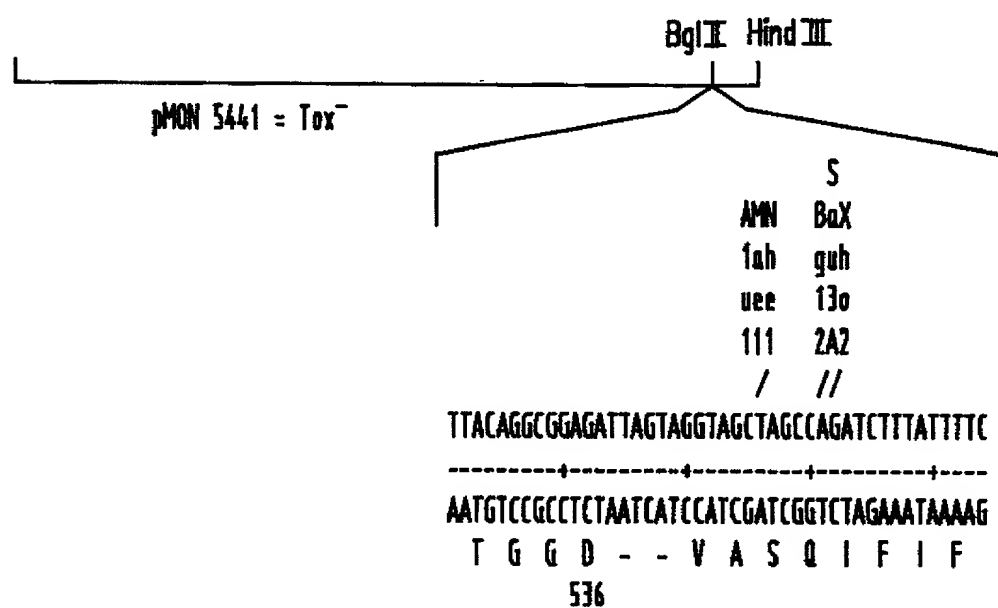
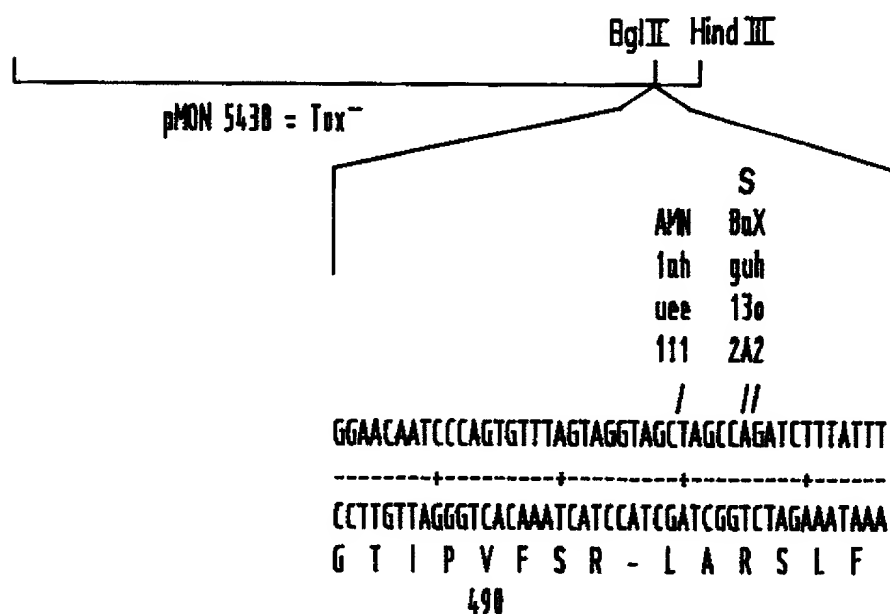
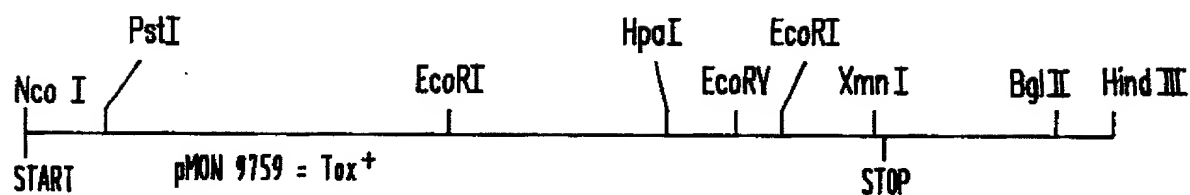
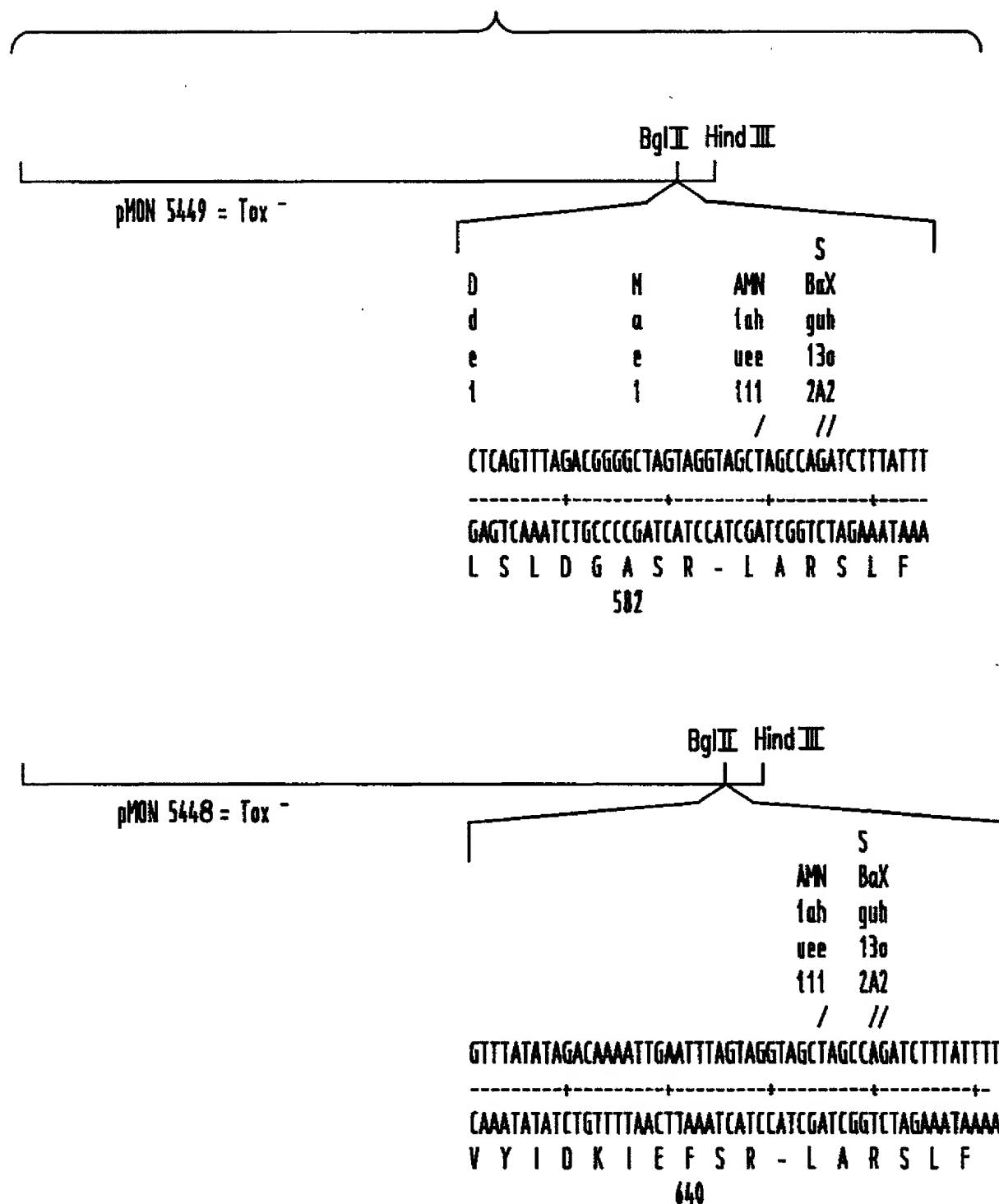
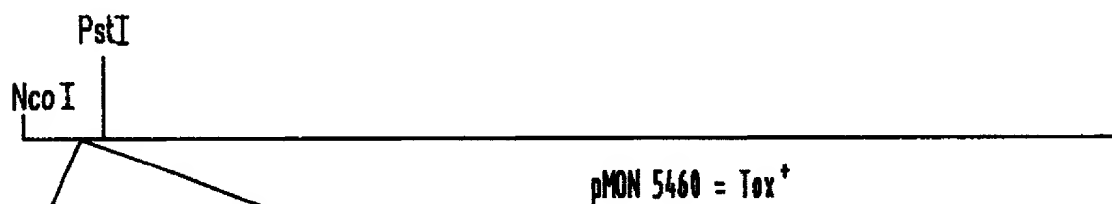
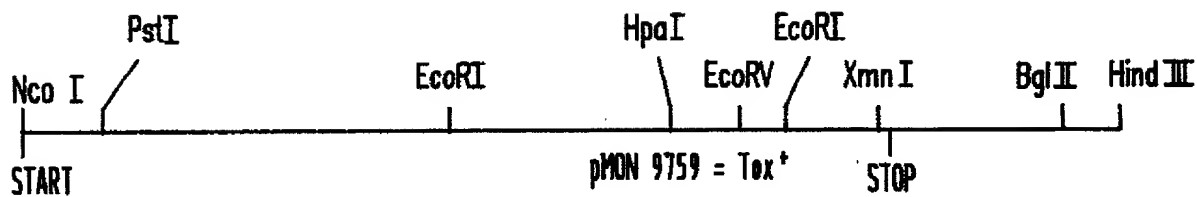


FIG.8A



THE INSERTS SHOW THE ACTUAL AMINO ACID  
SEQUENCE OF THE ALTERED B.t.t. PROTEINS.

FIG.8B



M	P
a	s
e	t
2	1

TATAAAGAGT TTTTAAGAATAACTGCAGATAATAATA

ATATTTCTCAAAAATTCTTATTGACGTCTATTATTAT

Y K E F L R I T A D N N T

48



NS	N M	F	M	A
ct	l b	o	a	l
ay	a o	k	e	u
11	3 2	1	1	1

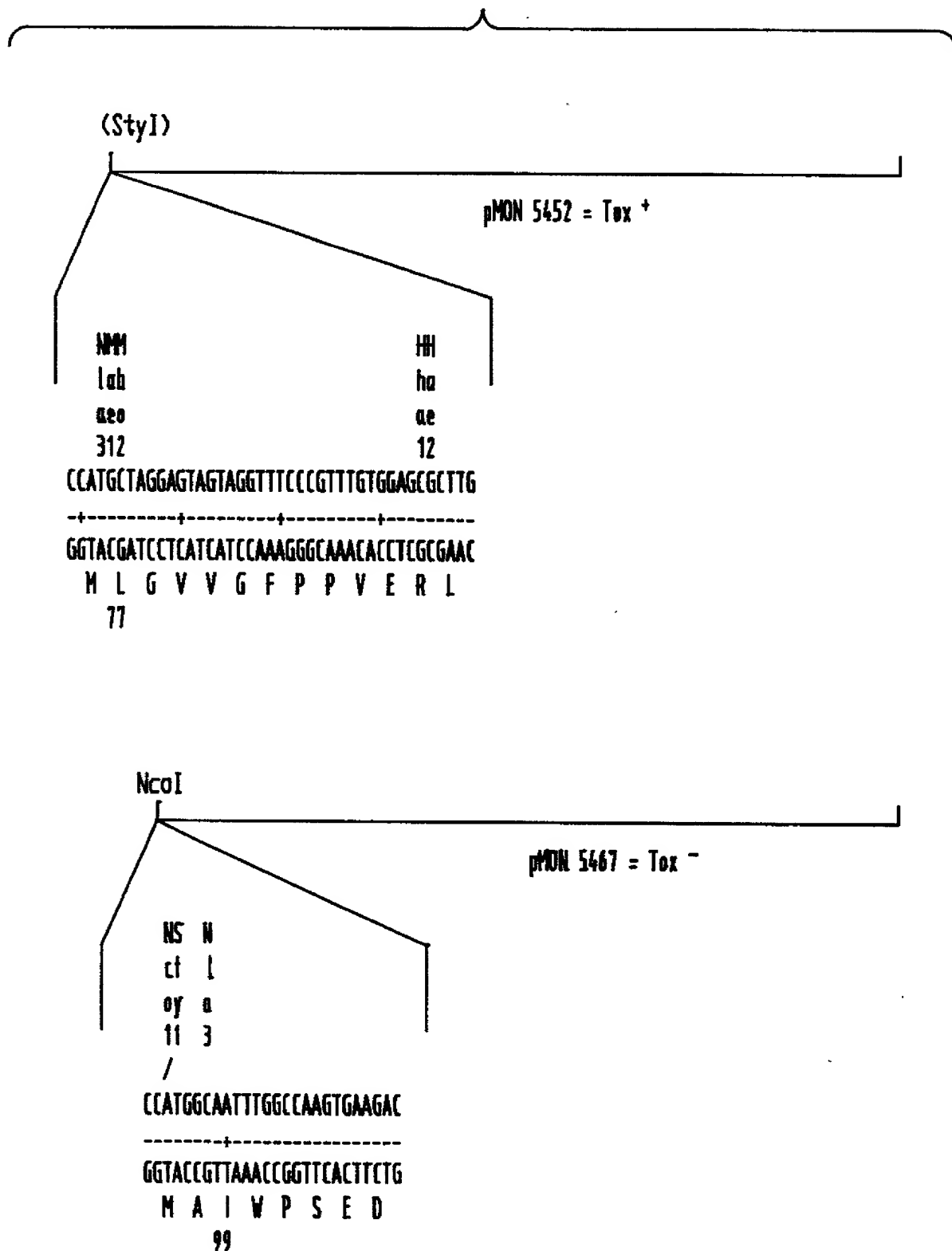
CCATGGATGCAGATAATAATACGGAGCACTAGATAGCTCT

GGTACCTACGTCTATTATTATGCCCTTCGTCTATCTGAGA

M D A D N N T E A L D S S

48

FIG.9A



THE INSERTS SHOW THE ACTUAL AMINO ACID  
SEQUENCE OF THE ALTERED B.t.f. PROTEINS.

FIG.9B

[illegible]

FIG. 10



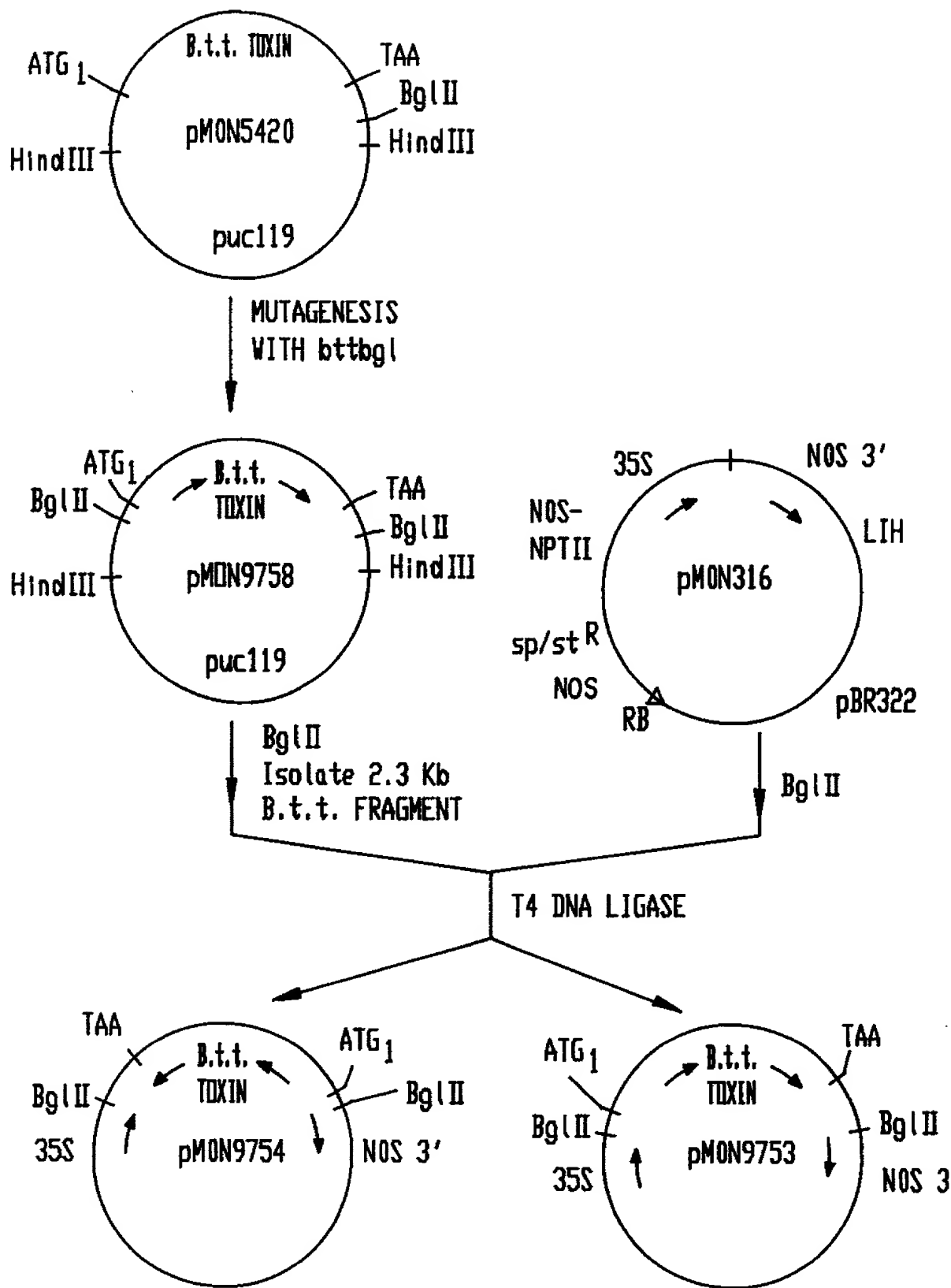


FIG.11

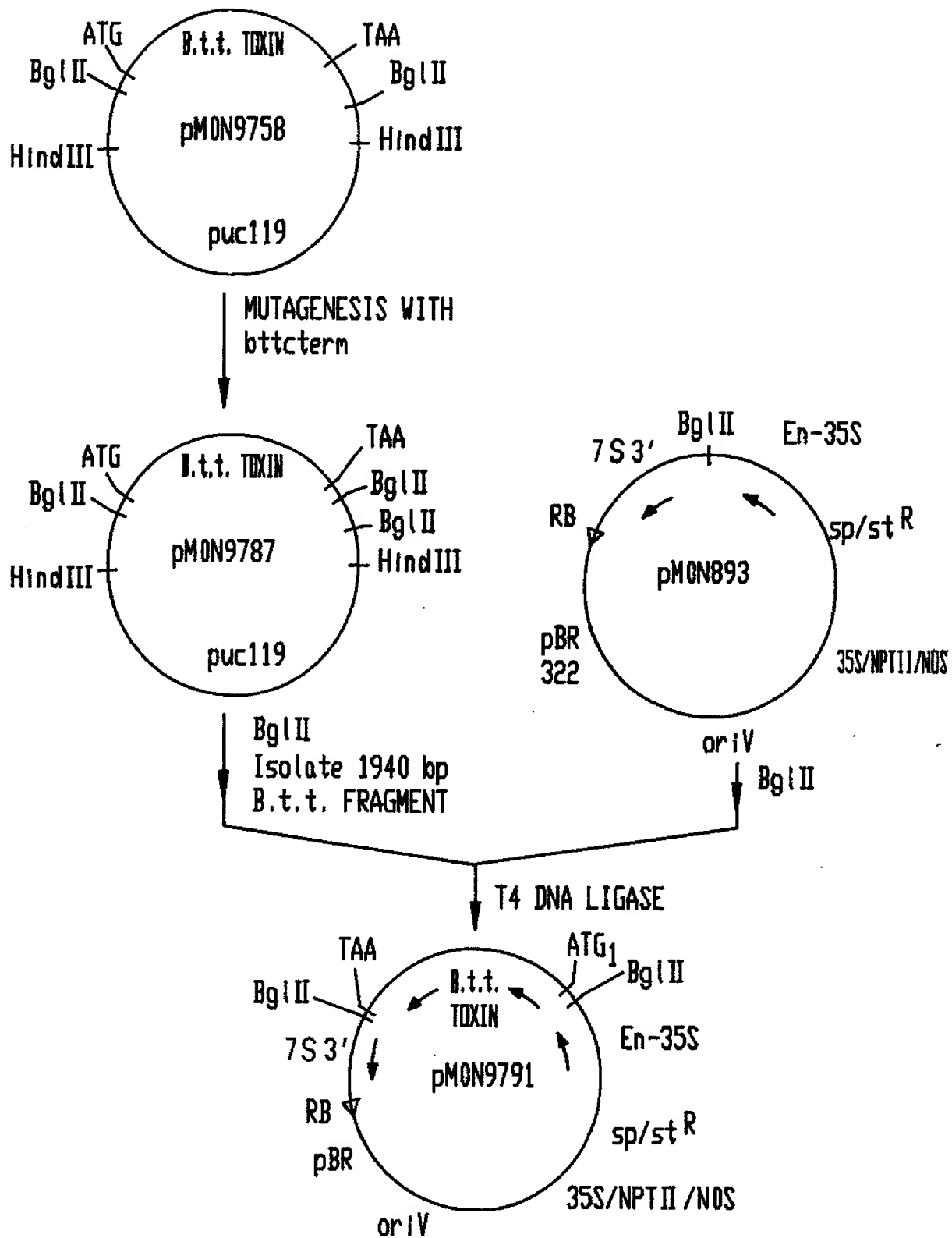


FIG.12

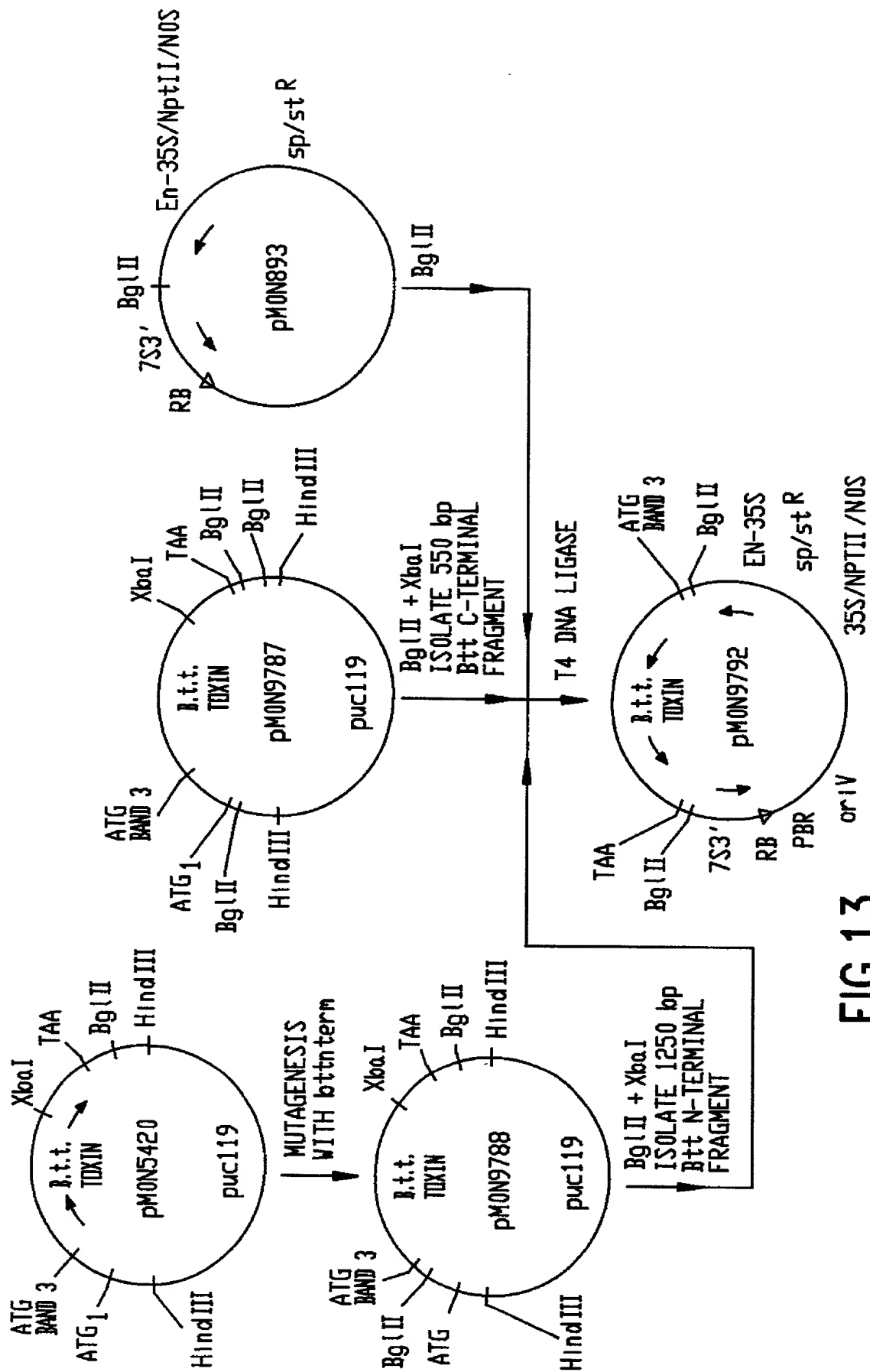


FIG.13

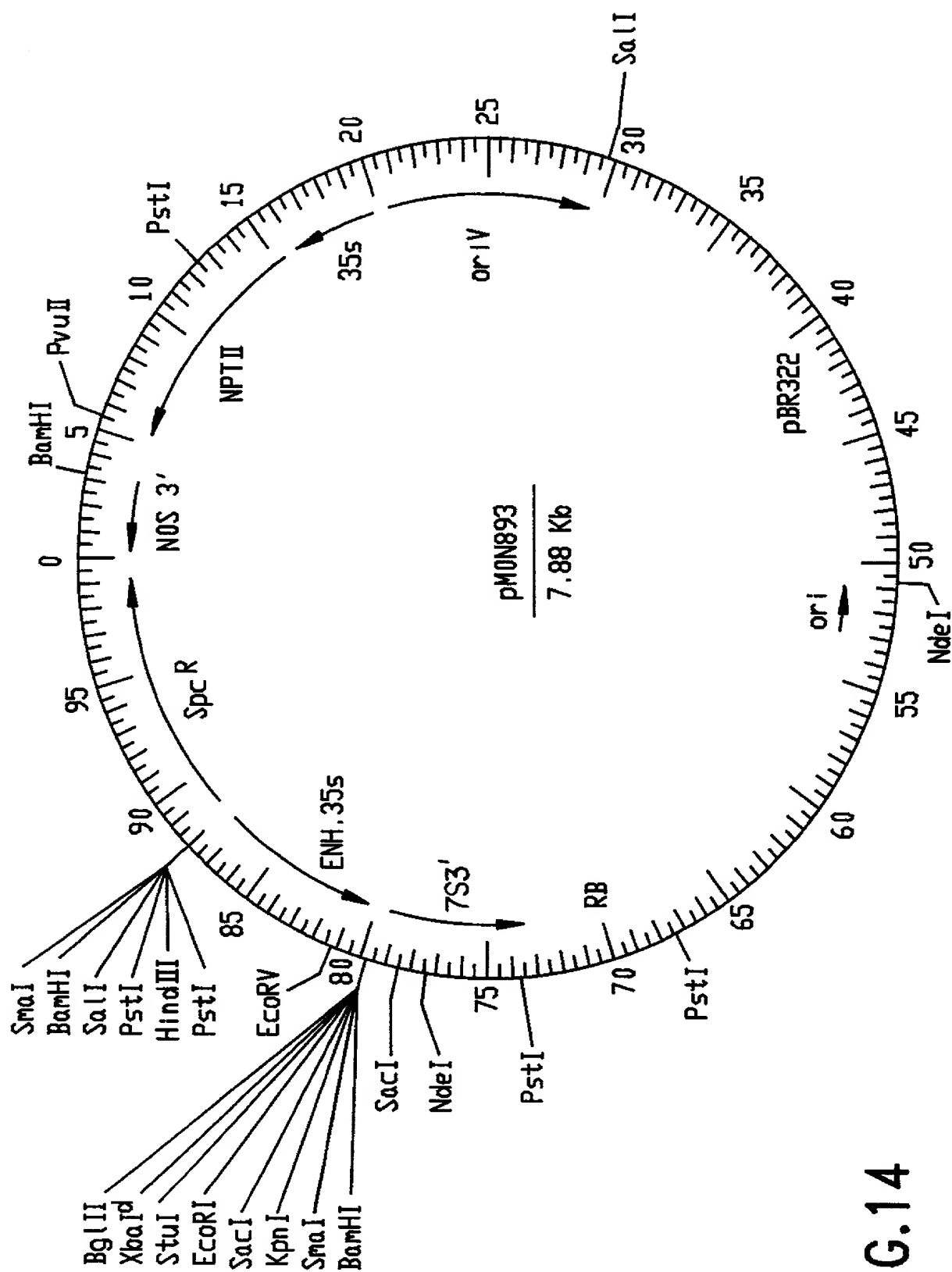


FIG. 14

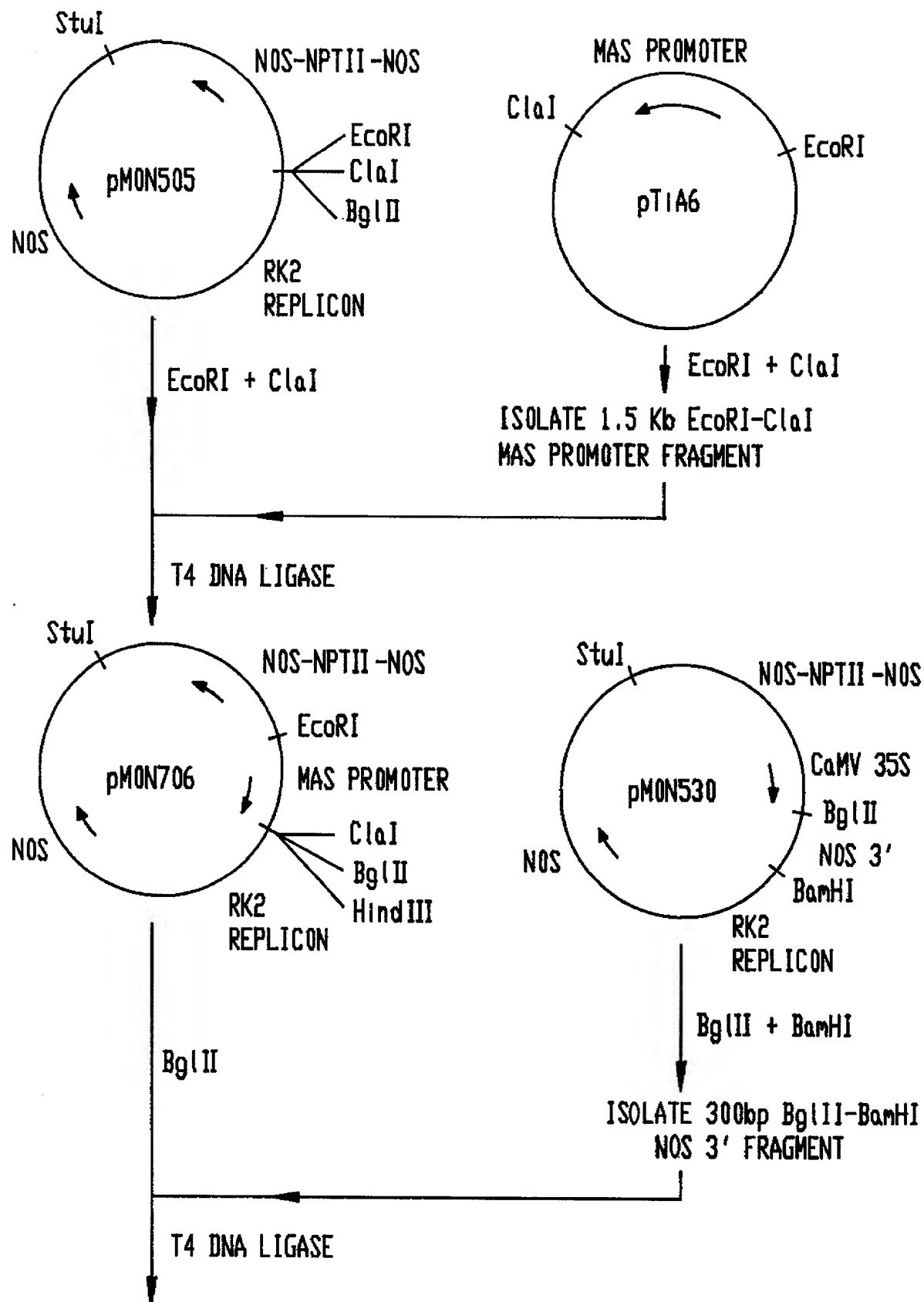


FIG.15A

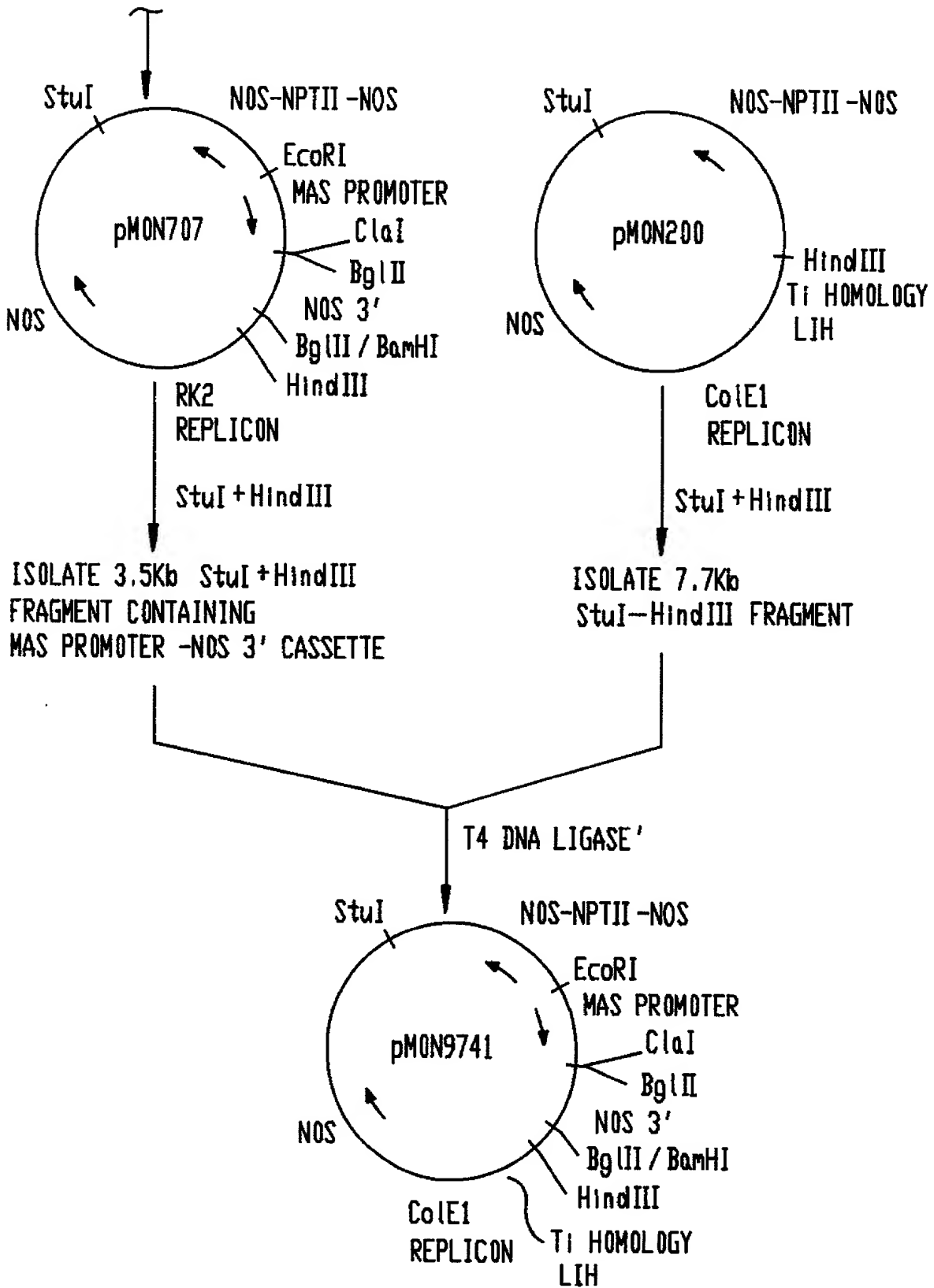
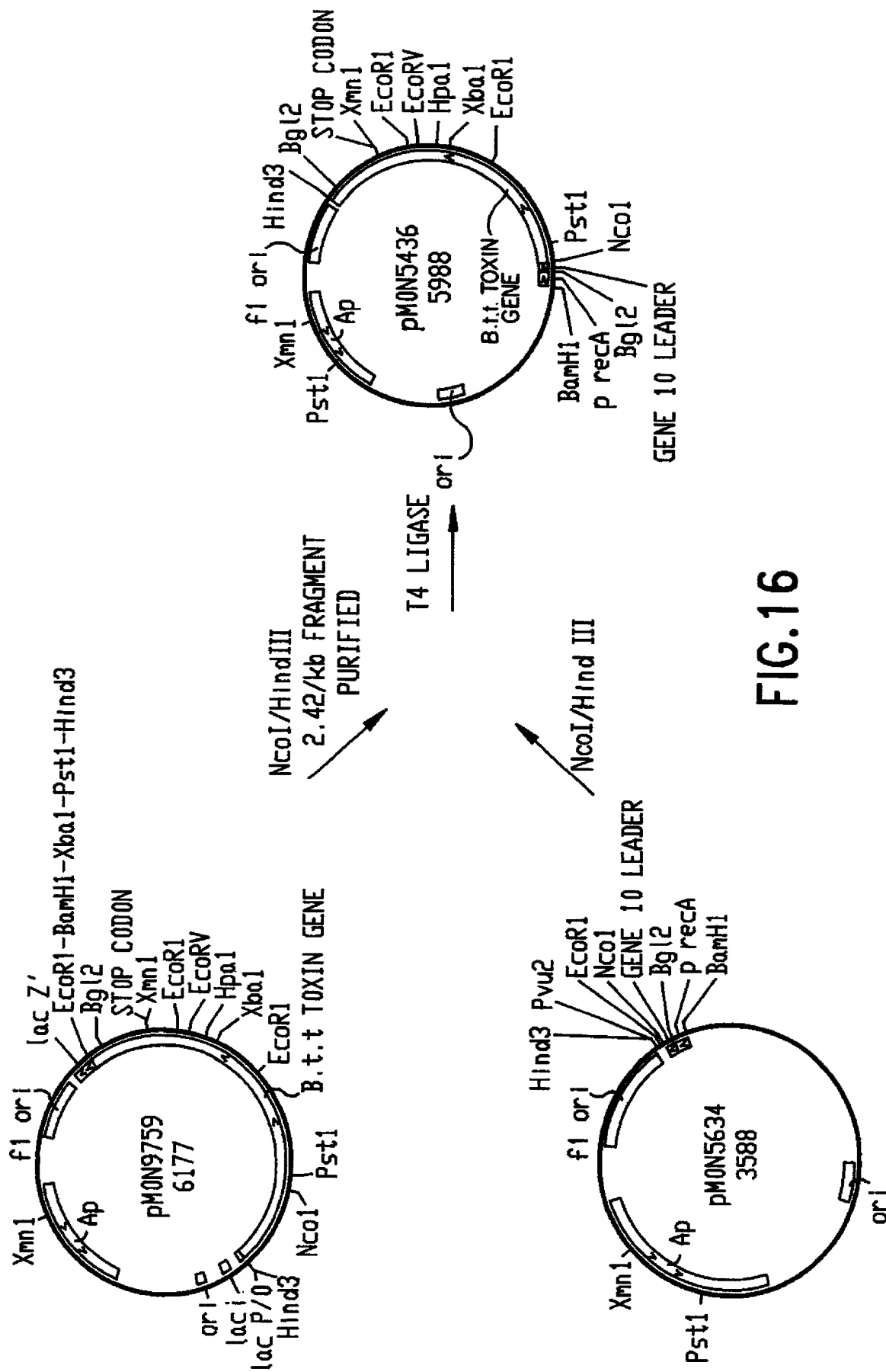


FIG.15B



T- DNA REGION IN AGROBACTERIUM ACO

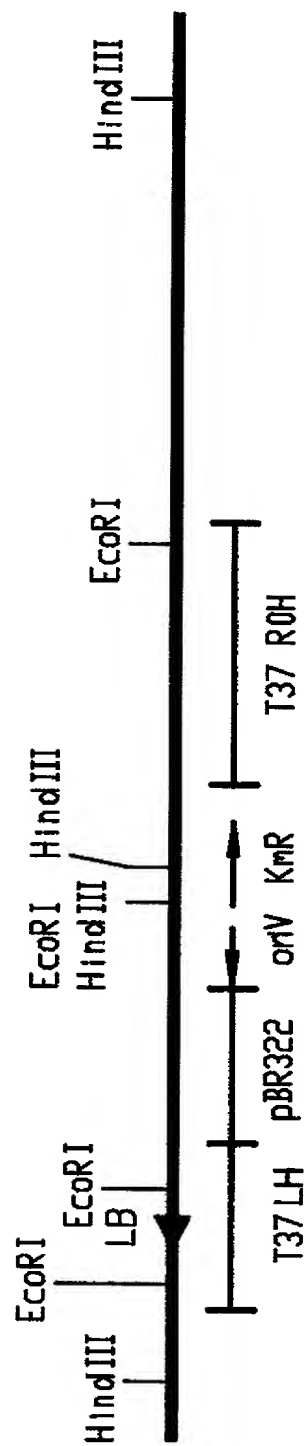


FIG. 17



DNA SEQUENCE FOR THE ENHANCED CaMV35S PROMOTER  
USED IN THE PREPARATION OF pMON893

\*

```

5' -AAGCTTGCAT GCCTGCAGGT CCGATG[GAG ACTTTTCAAC AAAGGGTAAT  50
    ATCCGGAAAC CTCCTCGGAT TCCATTGCCC AGCTATCTGT CACTTTATTG  100
    TGAAGATAGT GGAAAAGGAA GGTGGCTCCT ACAAATGCCA TCATTGGCAT  150
    AAAGGAAAGG CCATCGTTGA AGATGECTCT GCCGACAGTG GTCCCAAAGA  200
    TGGACCCCCA CCCACGAGGA GCATCGTGGA AAAAGAAGAC GTTCCAACCA  250
    CGTCTTCAAA GCAAGTGGAT TGATGTGATG]GTCCGATG[G AGACTTTTCA  300
    ACAAAGGGTA ATATCCGGAA ACCTCCTCGG ATTCCATTGC CCAGCTATCT  350
    GTCACTTTAT TGTGAAGATA GTGGAAAAGG AAGGTGGCTC CTACAAATGC  400
    CATCATTGCG ATAAAGGAAA GGCCATCGTT GAAGATGCCT CTGCCGACAG  450
    TGGTCCCAAA GATGGACCCC CACCCACGAG GAGCATCGTG GAAAAAGAAG  500
    ACGTTCCAAC CACGTCTTCA AAGCAAGTGG ATTGATGTGA ]ATCTCCACT  550
    GACGTAAGGG ATGACGCACA ATCCCACTAT CCTTCGCAAG ACCCTTCCTC  600
    TATATAAGGA AGTTCATTTC ATTTGGAGAG GACACGCTGA CAAGCTGACT  650
    CTAGCAGATC T - 3'                                     661
  
```

\* BRACKETED SEQUENCE INDICATED DUPLICATED ENHANCER SEQUENCE

FIG.18